DOCUMENT RESUME

ED 211 188 PS 012 531

AUTHOR Wedell-Monnig, Jacelyn; McNeil, Judy T.
TITLE National Evaluation of Head Start Educational Services and Basic Educational Skills Initiative. Project Report 4: Individual Testing Battery for Fall 1979.
INSTITUTION NTS Research Corp., Durham, N.C.
SPONS AGENCY Administration for Children, Youth, and Families (DHEW), Washington, D.C.
PUB DATE 18 Feb 80
CONTRACT HEW-105-78-1306
NOTE 27p.; For related documents, see PS 012 529-533 and TM 810 936-942. Tables may not reproduce clearly because of small print.

EDRS PRICE MF01/PC02 plus Postage.
DESCRIPTORS Auditory Perception; *Basic Skills; Cognitive Measurement; Communication Skills; *Delivery Systems; *Evaluation Criteria; *Evaluation Methods; Interpersonal Competence; Longitudinal Studies; *Measures (Individuals); Memory; *Outcomes of Education; Preschool Education; Psychomotor Skills; Visual Perception
IDENTIFIERS Basic School Skills Inventory; McCarthy Scales of Childrens Abilities; Preschool Interpersonal Problem Solving Test; *Project Head Start

ABSTRACT

The purpose and design of a longitudinal study of Project Head Start is summarized in this document, the fourth in a set of ten reports documenting the project. (Material from Project Report 3 was used extensively in preparing this document.) The study was designed to evaluate on a nationwide basis the Head Start Educational Services and to evaluate the Basic Educational Skills Program (BES), a recent demonstration project of Head Start. The intent of the longitudinal study was to follow children and their parents from preschool through fourth grade. The major evaluation topics addressed in the study included the following: the types of educational approaches used to teach Head Start participants (these approaches were compared to approaches used in alternative service settings); the effects of different intervention approaches on outcomes for Head Start participants; and a comparative analysis of Head Start participants in BES and Head Start participants not in BES. The dominant issues contained in this set of evaluation topics focused on the relationships between program delivery and child outcomes, including cognitive, communication and social skills. Cognitive and communication skills were measured by a battery of standardized tests, teacher reports, and observational techniques. Social skills were measured by a standardized test. The tests and how they were selected are identified. (Author/DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211188

NATIONAL EVALUATION OF HEAD START EDUCATIONAL SERVICES
AND BASIC EDUCATIONAL SKILLS INITIATIVE

Project Report 4: National Evaluation of Head Start Educational Services and Basic Educational Skills Initiative: Individual Testing Battery for Fall 1979

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
X Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

February 18. 1980

Contract No. HEW 105-78-1306

Authors: Jacelyn Wedell-Monnig, Ph.D.
Judy T. McNeil, Ph.D.

Project Director: Judy T. McNeil
NTS Research Corporation
2634 Chapel Hill Boulevard
Durham, NC 27707

Project Officer: Dennis Deloria

Prepared for:
The Administration for Children, Youth and Families
Department of Health, Education and Welfare
Office of Human Development

PS 012531

This report is one of ten submitted by NTS Research Corporation to the Administration for Children, Youth and Families under Contract No. HEW-105-78-1306.

Project Report 1: Design and Sampling

Project Report 2: Panel Recommendations

Project Report 3: Child Outcomes: Individual Testing Battery for the Fall 1979 Cohort

Project Report 4: National Evaluation of Head Start Educational Services and Basic Educational Skills Initiative: Individual Testing Battery for Fall 1979

Project Report 5: OMB Submission Package

Project Report 6: Description and Analysis of Program Data: Basic Educational Skills Initiative, 1979

Project Report 7: Description of Program Data: Head Start Sample Programs, 1979

Project Report 8: Generalizability Studies of the Child Test Battery and of the Alternative Classroom Environment Scales

Project Report 9: Procedures in the Implementation of the Study Design, 1978-79

Project Report 10: Phase II Summary Report: Fall 1979 Child Data Collection

TABLE OF CONTENTS

	PAGE
Introduction to the Evaluation	1
Identification of Skill Areas	5
Background of Evaluation	5
The Basic Educational Skills Initiative	5
Concept Papers	8
Identification of Skills Area for Testing	8
Screening of Tests	11
Test Battery	12
MSCA Subtests Selected	12
BSSI Subtests Selected	17
Preschool Interpersonal Problem Solving Test	19

TABLES

TABLE NUMBER		PAGE
1	Basic Educational Skills: Areas for Curriculum Focus	7
2	Areas of Skills to be Assessed	9
3	Tests Considered for Battery	13

Introduction to the Evaluation

Each year Head Start programs serve approximately 385,000 children and families, at a cost of over $700 million dollars--the largest annual federal expenditure for the education of preschool children. NTS Research Corporation is currently conducting the beginning phases of the largest and potentially most long-term evaluation of Head Start programs yet planned. The overall purpose of this evaluation is to assess the impact, on the children and parents being served, of: (1) the educational services component of Head Start programs nationwide, and (2) a new demonstration program within Head Start, Basic Educational Skills. The evaluation deals with a total of 65 Head Start programs and their families, potentially for 10 to 12 years.

Head Start is a nationwide program for preschool children of low-income families and is supported by the United States Department of Health and Human Services. Since its inception in 1965, the goal of Head Start has been to improve each child's ability to deal both with the present environment in which he/she lives and with later responsibilities in school and life. Head Start provides services not only in the educational area, but also in the areas of health, nutrition, parent involvement and social services. While this is an evaluation of the educational component, the study also gathers information on Head Starts' other components.

The Basic Educational Skills Program is the newest experimental project initiated by Head Start. As stated in the BES Program Guidelines:[1]

> The purpose of the project is to demonstrate and test effective educational strategies that help young children acquire developmentally appropriate educational skills in a climate that supports physical, intellectual, social, and emotional development. The project links Head Start programs and public schools, together with parents, in a collaborative effort to facilitate the child's learning experiences. The overall objective is to enhance school readiness, later school achievement, and to promote attitudes conducive to life long learning. The specific aim is to try out a variety of educational approaches to increase the child's basic educational skills, with emphasis on learning outcomes in oral language, reading, writing, mathematics, science, problem-solving, and learning attitudes. (Learning attitudes includes self-esteem as a learner, motivation, curiosity, initiative, purposefulness, persistence, task completion, and interest in life-long learning.) (p. 1)
>
> The goals of this project are to examine the effectiveness of a variety of approaches of implementing the four program elements: Curriculum, Parent Involvement, Teacher Attitudes and Behaviors, and Continuity. These program elements have been highlighted by reviews of research and educational practice as among the principal variables that affect children's educational achievement. The principal purpose of the BES field experiment is to determine the contribution of variables associated with the four elements to children's acquisition of basic educational skills. (p. 5)

Individual Head Start programs are operated by local grantees which are community action organizations, public school systems, tribal councils, and other non-profit agencies. Decisions about such issues as staff, curriculum, facilities, and recruitment procedures are made by a local policy council composed of parents and other members of the community. With this kind of local initiative, Head Start programs vary widely from one location to another in the ways in which they deliver services--within the guidelines of performance standards set by the Federal agency which funds and administers the program, the Administration for Children, Youth and Families (ACYF). The national office of

---

[1]Administration for Children, Youth, and Families, Basic Educational Skills Project: A Head Start Research and Demonstration Initiative in Collaboration With the Elementary Schools; Guidelines for Local Head Start Programs and Elementary Schools, March 1979.

ACYF also periodically funds special demonstration programs (such as BES) for the purpose of initiating testing, and disseminating innovative program ideas.

The regional offices of ACYF are responsible for monitoring local program service delivery to ensure that each program meets the required prformance standards. Evaluations of Head Start program outcomes are initiated by the national office of ACYF. The evaluations are diverse and include three distinct kinds:

- Evaluations of the basic National Head Start program, which use representative national samples of about 30 of the total 1,800 programs, and examine the areas of Health Services, Educational Services, Social Services, Parent Involvement, Services to Handicapped Children, Services to Migrant Children, and Staff Training.

- Evaluations of the Head Start demonstration programs, each of which involve about a dozen sites nationally, including, for example, Project Developmental Continuity, the Child and Family Resource Program, Home Start Training Centers, the Bilingual/ Bicultural Curriculum Project, and the BES Project.

- Projects to develop evaluation measures and methods; for example, the project to translate the ETS CIRCUS preschool test battery into Spanish, and the project to develop a comprehensive set of measures of program effects on preschool children.

This evaluation represents both the first and second categories above; it is a nationwide evaluation of the Educational Services of the basic Head Start Program and the evaluation of the newest Head Start demonstration project, Basic Educational Skills.

ACYF has actively gathered information over the past fifteen years regarding both program delivery and program effects. A unique aspect of the evaluation described in t is report is its focus on the relationship between program delivery and program effects on a national scale.

Overview of the Head Start Evaluation Design

This evaluation is a longitudinal study that will follow children and their parents from preschool through the fourth grade. The study focuses on the preschool and elementary experiences of three groups of children and their parents:

- o Those who participate in typical Head Start programs--the Head Start Study Group

- o Those who are eligible for Head Start but who are not enrolled--the Alternative Services (Control) Group

- o Those who participate in the Basic Educational Skills demonstration program--the BES Group

The major evaluation questions to be addressed in this study are as follows:

1. What different types of approaches do Head Start programs use to deliver educational services? How do these approaches differ in BES programs? How do these approaches differ from programs in alternative service settings?

2. How do different program approaches, and ways children and parents partake of the programs, affect the benefits achieved by children and parents?

3. Do children and parents in BES programs benefit more than children and parents in non-BES Head Start programs? Do children and parents in both BES and non-BES Head Start programs benefit more than children and parents in alternative service settings?

The dominant issues contained in this set of questions is a focus on the relationships between program delivery and program effects on a national scale.

As mentioned above, the evaluation is longitudinal and will study children (and their parents) from preschool through the fourth grade. Over the next three years, three cohorts (one small cohort and two full cohorts) of preschool age children will be studied. Each full cohort will be drawn from 16 Basic Educational Skills demonstration programs and from 30 randomly selected Head Start programs and their communities. It will involve 4560 children and their familier per cohort: 60 BES children from each BES program, 60 Head Start from each regular Head Start program, and 60 non-Head Start children from each regular Head Start program's community.

The study design is innovative for two reasons. First, it focuses on the relationships between process and outcome variables in preschool settings. Second, it is one of the first preschool field evaluations to incorporate a true experimental design. In those instances in which a participting Head Start program receives applications from more children than it can serve, random assignment of children into the Head Start program and a non-Head Start group will be implemented. The use of this study plan will, hopefully, result in one of the strongest research efforts ever conducted under the auspices of ACYF.

The impact of the educational services component and the Basic Educational Skills program on children will be measured in multiple ways, including classroom observations (piloted during Spring 1979 site visits), individual child in-classroom observations (the classroom observation system will be modified for this purpose), academic performance in kindergarten and beyond, possibly observations of child and parent together, and, of course, individual testing during the Head Start year. This report concerns the history and development of the first individual testing battery.

The RFP for this evaluation called for child outcomes to be measured in four areas:

- "Developmentally appropriate pre-academic and cognitive skills as measured by standardized test and teacher reports of classroom behavior and potential. Outcomes to be assessed should parallel the following domains: language, reading, writing, mathematics, science, and problem-solving.
- Motivational orientation toward school experiences, toward learning, self-concept as learner, and ability to delay gratification.
- Indications of sociability toward peers, toward teachers, and toward adults in general.
- Long term staying power of the experience (ability to adapt to later schools, assignments to special education classes, patterns of grade retention)."

Emphasis in this present battery will be on the first point, that is, cognitive and communication skills. Motivational orientation and self-concept are difficult to accurately measure in the young child. These areas will later be assessed primarily by observational measures. The third area of social skills is tapped by one test in the battery. Finally, "long-term staying power" will clearly be examined later by means of observations, individual testing and school records.

In October, 1979, a draft of "Project Report 3: Child Outcomes: Individual Testing Battery for the Fall 1979 Cohort" (Wedell-Monnig and McNeil) was submitted to the Administration for Children, Youth and Families. Parts of that report have been exerpted to form the following paper.

Identification of Skill Areas

Background of the Evaluation

This evaluation is twofold: it will assess the impact of the educational services component of Head Start on young children by comparing children who have experienced Head Start with a group of children who do not participate in Head Start. Secondly, the effectiveness of the Head Start Basic Educational Skills project is being studied. One battery has been designed for use in both aspects of the evaluation. The BES Phase I quidelines (Klein, 1978) thus were used heavily in identifying areas in which to test children.

Basic Educational Skills Initiative

The ACYF Basic Educational Skills project was initiated in response to President Carter's "call for efforts to assist children in acquiring the basic skills they need to function successfully in a complex society...although these [skills] will contribute to children's later academic learning, they also include skills in communication and problem solving as well as positive learning attitudes." (Klein, 1978, p.1)

What are basic educational skills? "Within the context of Head Start and its goals, developmentally appropriate skills are those which are necessary for the later acquisition of skills and learning commonly recognized as reading, writing, and mathematics, but which also include language, science, problem solving and the development of positive attitudes toward learning." (op.cit., p.2) These "preskills" are not, it is pointed out, watered down versions of the academic skills usually taught in the primary grades. This is a vital point:

there is substantial evidence in the cognitive development literature that the cognitive processes of four or five year olds and seven or eight year olds differ qualitatively, not just quantitatively. According to Piaget, for example, (see Flavell, 1963, for a review of the theory), the younger child is in the stage of preoperations, characterized by concrete and static thought, inability to take another's point of view, and an inability to take into account more than one aspect of a problem at a time. The older child is in the sub-stage of concrete operations. The concrete operational child is better able to deal with abstractions, can take points of view other than his own, and can understand that two aspects of a problem may compensate for each other. One cannot simply follow a continuum of skills from age 4 to age 7; differences in thought processes must be taken into account. The task of identifying appropriate pre- or basic skills is thus not a simple one.

Some of these skills are identified in the BES guidelines: "Continued interest in learning, the ability to understand and use language in speaking and listening to others, beginning to understand concrete relationships such as quantity, size and shape, to discriminate between different shapes and colors, and to begin to recognize symbols such as letters and words." (Klein, 1978, p.3)

The BES Phase I guidelines paper (Klein, 1978) and the Abt approach paper (Bache, Brush, Larson and Mathis, 1978) provided us with an outline to the definition of the basic skills to be considered in designing the battery. The guidelines in the Abt paper delineate the BES curriculum into five basic areas: Oral Language, Reading, Writing, Mathematics, and Science/Problem Solving. These five areas are presented in Table 1 along with examples of teaching-learning activities that might be used in the Head Start year.

Areas for Curriculum Focus

Oral Language	Reading	Writing	Mathematics	Science Problem Solving
NARRATION o Discussing events and asking questions in the context of conversation	COMPREHENSION o Listening to stories and engaging in conversations about them	MOTOR COORDINATION o Using a variety of art materials for painting, drawing	CLASSIFICATION o Grouping and ordering sizes and shapes in the context of block play	EXPERIMENTATION o Mixing paint or playdough for use in art activities
COMMUNICATION o Engaging in dialog in dramatic play	SYMBOLIZATION o Pretending to read by using pictures and props symbolically in dramatic play	SHAPE DISCRIMINATION o Recognizing two dimensional shapes, as a form in pictures or one's name	NUMERATION o Matching shapes, colors, or symbols to identical items during games with manipulatives	EXPLORATION o Balancing structures or creating and using inclines in block play
DESCRIPTION o Describing block buildings, specific block shapes and accessories	SEQUENCING o Using manipulative materials to sequence pictures or arrange page	SENSORY INTEGRATION o Skipping, hopping, and climbing using alternate feet/legs, hands/arms in outdoor play	QUANTIFICATION o Measuring ingredients for cooking activities	COMPARISON o Using a magnifier or balance with a variety of manipulative materials
AUDITORY DISCRIMINATION o Differentiating sounds in the context of signing or listening to music	FIGURE-GROUND DISCRIMINATION o Looking at pictures in books with attention to shapes, colors, and outlines		NUMERICAL RELATION o Clapping, foot-stomping, and beating instruments to music	INFERENCE o Baking or cooking to anticipate foods' changes in form, color, taste, and smell
CONCEPTUALATIZATION o Taking field trips to familiar places in the community, such as grocery, fire station	ORIENTATION o Taking field trips to places that make written language meaningful, e.g., library		ESTIMATION o Running, throwing, and catching a ball in outdoor play	OBSERVATION o Taking field trips or short walks to observe nature

Concept Papers

Although the BES guidelines are fairly complete in presenting examples of classroom activities that relate to the BES areas, identifying testable precursor or basic skills is somewhat more difficult. Therefore, NTS took the further step of consulting several child development experts for their recommendations. We decided to ask for "concept papers": short, concrete reviews of current thinking and precursors to basic academic skills. The concept papers focused on the areas of reading/writing and mathematics. We felt that NTS had the in-house competence to determine precursors in the language area. In place of a concept paper for the science/problem-solving area, we arranged for a one-day consultation.

Robert Calfee of Stanford University contributed the paper on assessment of precursor skills to reading and writing. Joan Heller and Lauren Resnick wrote the paper on "Skills and Concepts Preparatory to Formal Schooling and Mathematics." Finally, Jacelyn Wedell-Monnig of NTS met with Irving Siegel of ETS to discuss skills essential to the science/problem-solving area.

Identification of Skills Areas for Testing

The next step in selecting a battery was to determine a philosophy by which to select skill areas to assess. We recognize the ACYF/Abt BES areas as guidelines for curriculum development. Overlap between the areas is so great, however, that it is not entirely appropriate to use these areas for test selection. Our approach was to first closely examine each of the BES areas in order to identify independent elements. Secondly, we studied the concept papers and selected some additional elements. The result was the list of areas and elements presented in Table 2. Six skill areas were selected, within which prospective tests were examined.

TABLE 2

Areas of Skills to be Assessed

Factual Knowledge

Letter Names
Number Names
Color Names
Shape Names
World Knowledge

Memory Skills

Visual Memory
Auditory Memory

Cognitive Skills

Spatial Reasoning
Classification
Numeration and quantification
Problem Solving

Communication Skills

Oral Language Production
Vocabulary
Creativity
Syntax

Oral Language Comprehension
Vocabulary
Comparatives

Perceptual/Psychomotor Skills

Visual Discrimination
Auditory Discrimination
Fine Motor Skills
Gross Motor Skills

Social Skills

Both the reading/writing and mathematics concepts papers emphasize the need for the child to master certain skills before the later respective skills can be learned in the primary grades. Many of these skills concern the acquisition of facts such as letter, number, color, and shape names. We included world knowledge, since the child who is not in command of general and basic facts will be at a certain disadvantage in acquiring further information. Likewise, the child whose visual and auditory memory skills are not well developed will not be able to master basic reading/writing and mathematics concepts.

Cognitive skills are essential in gaining understanding of number systems as well as in solving problems.

Communication skills are basic to the areas of language and science/problem solving, as well as to reading/writing. Finally, as pointed out by Heller and Resnick in their concept paper, certain perceptual/psychomotor skills are essential to success in many primary grade tasks.

Social/emotional development is not emphasized in this list to the same extent as cognitive skills. We feel that such development cannot be fully assessed by means other than direct observation of the child. High/Scope came to the same conclusion in their recommendations for measures for Project Developmental Continuity (Love, Wacker, and Meece, 1975). In terms of self-concept, for example, High/Scope stated that:

> The fundamental difficulty with the construct of self-concept is the multiplicity of definitions. Many critics (Coller, 1971: Crowne and Stephens, 1961; Taylor, Winne, and Marx, 1975) have concluded that what are called self-concept measures actually assess many different constructs. No agreed upon operational definition of self-concept exists, largely because there is no adequate theory within which it can be based. As a result there are a great variety of instruments and techniques for narrowly defined aspects of self-concept, so that "it is likely to be defined as 'that which a self-concept test measures'" (Walker, 1973). This, of course, makes the selection of any one instrument as a measure of program effectiveness especially problematic. (Love, Wacker, and Meece, 1975, p.18).

Screening of Tests

A summary form was developed for screening each test which was examined for possible inclusion in the battery. After forms on all considered tests were complete, the forms were studied to determine which tests met the following criteria (order below does not indicate relative importance):

1. Yields relevant and analyzable data. The test should be relevant to the skills we have determined are important to assess. The results must be interpretable in light of Head Start program objectives.

2. Free of cultural biases to the extent possible. Many cultural groups are represented in the population under study. We rejected many tests on the basis of being appropriate for only a few of these groups. Tests must be fair both in terms of content and scoring of responses.

3. Normed (or at least used) on a population relevant in terms of social class, cultural groups, and age. Although we do not plan to compare our data to existing norms, previous use of the test with a similar population gives indication of whether other researchers found the measure to be reasonably relevant to that population.

4. Reliability and validity data available and acceptable. Preferably used in other large scale evaluations of preschool and elementary level programs to improve the comparability of results from one evaluation to another.

5. Instrument must have high generalizability (see McNeil, et al, 1978, and Project Report 8 for a full discussion of generalizability).

6. Instrument can be readministered to same children (repeatability). As this evaluation involves pre- and post-testing, we must be able to use the same tests at least twice for each child.

7. Able to be administered by trained paraprofessionals. In most sites, children will be tested by local data collectors, hired and trained by NTS. It is vital that these people can master the test in a short but intensive training session. This means that testing procedures must be very clear and precise.

8. Able to be administered quickly and with minimal equipment. The testing time per child will be no more than one hour. Therefore, each test in the battery must be brief and concise. Minimal equipment needs reduce cost and difficulty of testing.

9. Individually administered. All tests will be given in individual sessions with only the tester and child present, if at all possible.

10. Child should be unaware that she/he has failed an item. Tests were carefully examined to determine if a child could become unnecessarily discouraged. Feedback on selected tests is neutral as often as possible to avoid giving the child information on his or her performance. We feel that neutral feedback will have the least effect on the child's performance on subsequent tests.

Many tests used in other assessments were not even screened as they did not meet more than one or two of the criteria. The tests that we screened are listed in Table 3. Also in Table 3 are brief summaries of the data we examined to make preliminary test selections.

Two assessment batteries and one test emerged from the screening as highly relevant to our evaluation: The McCarthy Scales of Children's Abilities (MSCA) (McCarthy, 1972), the Basic School Skills Inventory (BSSI) (Goodman and Hammill, 1975), and the Preschool Interpersonal Problem Solving task (PIPS) (Spivak and Shure, 1974). Neither of the two batteries were judged to be appropriate for use in its entirety. The MSCA is a lengthy test and includes many subtests not of interest to us. The BSSI not only contains material not useful to us, but was designed as an informal assessment device to be used by teachers and needed to be redesigned in order to be administered in a consistent manner.

Certain subtests of the MSCA were used by High/Scope in the assessment of Project Developmental Continuity. High/Scope's decisions on the subtests for their own batteries were taken into account in designing the present battery.

The specific MSCA and BSSI subtests selected and the modifications made are discussed in the next session, along with a summary of the PIPS.

Test Battery

MSCA Subtests Selected

The MSCA is a well known, though relatively recent (1972) individual assessment measure for two-and-a-half to eight-and-a-half year old children. The MSCA was standardized on a representative sample of 1,032 children (stratified according to age, sex, color, geographic region, father's occupation and

TABLE 3

Tests Considered for Battery

Variables / Test	Contents						Age Range	Length of Test	Cultural Biases	Psychometric Variables: Relevant Norming Population	Psychometric Variables: Validity Data	Psychometric Variables: Reliability Data	Comments
Animal Crackers						X	Preschool Grade 1	20-45 mins.	Not recommended for bilinguals	No (largely white)	Data from Gumgookies. Sig. correlations with teacher ratings of motivation	r = .94-.98 (KR-20)	Many references to school and teacher - bias against control group.
Basic School Skills Inventory: Basic Information	X						4-6 years	5-7 mins.	Some items	Yes	High correlations with MRT and teacher ratings (entire battery)	r = .60-.76 (Internal consistency coefficients) High test-retest correlations with small n	Intended to be administered by teacher.
Basic School Skills Inventory Number Readiness	X		X	X			4-6 years	7-10 mins.	No	Yes	High correlations with MRT and teacher ratings (entire battery)	r = .78-.90 (Internal consistency coefficients) High test-retest correlations with small n	Intended to be administered by teacher.
Basic School Skills Inventory: Reading Readiness	X		X		X		4-6 years	7-10 mins.	Auditory discrimination items	Yes	High correlations with MRT and teacher ratings (entire battery)	r = .86-.98 (Internal consistency coefficients) High test-retest corelations	Intended to be administered by teacher
Boehm List of Basic Concepts	X			X			Grades K-1	15-20 mins.	?	Insufficient data	No data reported	r = .88-.90 (split half)	Concepts may be too advanced for 4 year olds
CIRCUS: Copy What You See (A)					X		preschool	30 mins.	No	No (largely white, middle class)	No studies reported	r = .87-.90 (alpha) r = .89-.90 (split half)	Cannot be machine scored
CIRCUS: Do You Know (A)	X						preschool	30 mins.	Definite biases	No (largely white, middle class)	No studies reported	r = .77-.79 (alpha) r = .79-.80 (split half)	Very biased in favor of white middle class
CIRCUS: Finding Letters and Numbers (A)	X				X		preschool	30 mins.	No	No (largely white, middle class)	No studies reported	r = .86 (alpha) r = .86-.87 (split half)	
CIRCUS: How Much and How Many (A)	X		X	X			preschool	30 mins.	Biased	No (largely white, middle class)	No studies reported	r = .86-.87 (alpha) r = .87-.88 (split half)	Quantitative concepts measure may be confounded with vocabulary comprehension
CIRCUS: How Things Work (A)				X			preschool	30 mins.	Biased	No (largely white, middle class)	No Studies reported	r = .43-.78 (alpha) r = .75-.81 (split half)	Understanding of factors; scores difficult to interpret
CIRCUS: Listen to the Story (A)				X			preschool	30 mins.	Biased	No (largely white, middle class)	No Studies reported	r = .60-.79 (alpha) r = .79-.80 (split half)	Receptive language

TABLE 3 (Continued)

Tests Considered for Battery

Variables / Test	Contents						Age Range	Length of Test	Cultural Biases	Psychometric Variables: Relevant Norming Population	Psychometric Variables: Validity Data	Psychometric Variables: Reliability Data	Comments
CIRCUS: Look Alikes (A)					X		preschool	30 mins.	No	No (largely white, middle class)	No studies reported	r = .52-.84 (alpha) r = .86-.88 (split half)	Measures ability to mark letters, number, shapes, objects
CIRCUS: Say and Tell (A)					X		preschool	untimed	No	No (largely white, middle class)	No studies reported	r = .49-.72 (alpha)	Productive language. May be difficult to machine score
CIRCUS: Think it Through (A)			X				preschool	30 mins.	Relatively unbiased	No (largely white, middle class)	No studies reported	r = .81-.82 (alpha) r = .81-.82 (split half)	problem solving; classification, sequences
Cognitive Abilities Test (Primary Battery)			X				K - Grade 2	Untimed; averages 54 mins.	No	Data not yet available	Data not yet available	For previous form - r = .92-.94 (KR-20)	size, position, quantity, classification, vocabulary, relational concepts
Cooperative Preschool Inventory			X	X	X	X	3-6 years	15 mins.	Intentional see Comments	Yes	r = .39-.65 with Stanford Binet IQ	r = .88-.91 (KR-20) r = .84-.92 (split half)	Designed to demonstrate effects of intervention and degree of disadvantage
McCarthy Scales of Children's Abilities (MSCA): Block Building			X		X		2 1/2-8 1/2 years	3-5 mins.	No	Yes	High position correlation (.70-.85) of total score (entire battery) with WPPSI and Stanford Binet IQ's. Moderate correlations (.30-.49) with first grade achievement tests. Very small sample (n=35)	Average split half coefficients: General cognitive scale, r = .93 Perceptual-Performance scale, r = .84 Verbal scale, r = .88, Memory scale, r = .79 Quantitative Scale, r = .81 Motor Scale r = .79	Scoring would have to be revised; too precise is is for real time scoring.
MSCA: Puzzle Solving			X					5-10 mins.	No				Consistent scoring may be a problem
MSCA: Pictorial Memory		X		X				3-5 mins.	Yes some items				Would be appropriate if better items selected.
MSCA: Word Knowledge				X				3-5 mins.	Yes some items				Stimuli appear inappropriate for use with this population
MSCA: Number Questions			X					3-5 mins.	No				Too difficult for this population
MSCA: Tapping Sequence		X	X					3-5 mins.	No				Too difficult for this population
MSCA: Verbal Memory		X		X			4-8 1/2 years	3-5 mins.	No				Used in POC evaluation
MSCA: Leg Coordination					X		4-6 1/2 years	3-5 mins.	No				
MSCA: Arm Coordination					X		2 1/2-8 1/2 years	3-5 mins.	No				Some items difficult for examiner.

TABLE 3 (Continued)

Tests Considered for Battery

Variables / Test	Contents						Age Range	Length of Test	Cultural Biases	Psychometric Variables: Relevant Norming Population	Psychometric Variables: Validity Data	Psychometric Variables: Reliability Data	Comments
MSCA: Imitative Action					X		2 1/2 - 8 1/2 years	3-5 mins.	No	Yes		Test-retest coefficients General Cognitive scale, r = .89-.91 Perceptual Performance scale, r = .78-.84 Verbal scale, r = .82-.89 Memory scale r = .79-.93 Quantitive scale, r = .89-.90 Motor scale, r = .69-.78	
MSCA: Draw A Design			X		X			5 mins.					Scoring may be difficult
MSCA: Draw A Child			X		X			3-5 mins.	No				Scoring may be difficult
MSCA: Numerical Memory		X					4- 6 1/2 years	3-5 mins.	No				Recal Digits
MSCA: Verbal Fluency			X	X				1-2 mins.	Some categories				Modified and used in PDC evaluation
MSCA: Counting and Sorting	X	X					2 1/2 - 8 1/2 years	3-5 mins.	No				
MSCA: Opposite Analogies			X	X				3-5 mins.	No				Difficult to translate into Spanish
MSCA: Conceptual Grouping	X		X	X			4- 6 1/2 years	3-5 mins.	No				
Peabody Picture Vocabulary Test				X			3.3 - 17.6 years	Average 10-15 minutes	Somewhat to white, middle-class, sexist	no; normed in one city	Predictive validity-PPVT scores at begining of 1st grade correlated with MAT scores at end of year word knowledge r = .35 reading r = .39	r = .67 for 6 year olds (Form A and B)	
Preschool Embedded Figures Test					X		3-5 years	25 mins.			Correlates .55-.67 with WPPSI Block Design subtest	r = .74-.91 (split half) r = .69-.75 (test-retest)	Visual discrimination
Preschool Interpersonal Problem Solving Test						X	4-5 years	20-25 mins.	No; multiethnic pictures included	No; normed on inner city only	Sig. correlations with teacher rating of social adjustment	r = .86-.89 (interrater) r = .59 (test-retest)	Recommend using only first half of test; used in PDC eval.
SIMPA: Bender Visual Motor Gestalt Test					X		5-11 years	20 mins.	No	No Data	No Data	No Data	More appropriate for neurological assessments
SIMPA: Physical Dexterity Tasks					X		5-11 years	20 mins.	No	No Data	No Data	No Data	More appropriate for screening for abnormalities
Stutnback Test of Reading Readiness	X	X	X	X	X		Grades K-1	40 mins.	Yes	No	?	r = .58-.94 (KR-20) r = .95 (split half)	Many items outdated
Wechsler Preschool Primary Scale of Intelligence Block Design			X		X		4- 6 1/2 years	15-20 mins.	No	Yes	Not reported	r = .76-.88 (split half)	Used in PDC evaluation

and urban-rural residence), with the sample closely matching the 1970 census data.

Kaufman and Kaufman (1973) found evidence that "performance on the cognitive scales in the MSCA is not a function of color" (page 205). Using the standardization sample, Kaufman and Kaufman identified 148 black-white pairs, matched on age, sex, and father's occupation. The sample was grouped into three age categories: 2-1/2 to 3-1/2, 4 to 5-1/2, and 6-1/2 to 8-1/2. Few significant differences were observed in performance, particularly in the subtests selected for use in the present evaluation. The specific subtests selected are described below:

Block Building. The child is asked to duplicate four simple structures built by the tester out of one-inch cubes. Performance is untimed and even partial success earns some credit. Block Building is used to evaluate spatial reasoning, visual discrimination, and fine motor dexterity.

Pictorial Memory. The child is shown a card which has six colored pictures of familiar objects. The tester shows the card for ten seconds, while naming the objects. The card is then removed and the child tries to recall the objects. We modified the original list of items, as we felt that some of the original objects (specifically, padlock and paperclip) were inappropriate for the present population. Pictorial Memory measures the child's visual memory and productive vocabulary skills.

Verbal Memory. The part of this test we utilized contains graded series of words and sentences to be repeated by the child. The first two items contain concrete concepts likely to be within the child's receptive vocabulary; the next two items contain two-syllable words which are more abstract in meaning and the last two items are full sentences. Verbal Memory is used to test several skills, including auditory memory and productive and receptive vocabulary.

Draw-a-child. Boys are asked to draw a boy and girls are asked to draw a girl. Human figure drawings are frequently used in clinical assessment as they provide information about the child's level of cognitive functioning as well as his personality. In this evaluation, the drawings are also useful in assessing perceptual motor development and fine motor coordination.

Verbal Fluency. The child's productive vocabulary and ability to classify and to think categorically are measured by asking the child to think quickly of words falling into each of four categories.

Conceptual Grouping. The child's ability to classify and to generalize is tested with the aid of a set of twelve blocks in different shapes, sizes, and colors. Conceptual Grouping also assesses the child's factual knowledge of color and shape names. Nine problems are presented orally, with the child responding by appropriate selection, grouping, or elimination of the blocks.

BSSI Subtests Selected

The MSCA subtests were selected, not only for their appropriateness for the present evaluation, but also for the excellent statistical data available. The data available on the Basic School Skills Inventories are much more limited and less impressive. Nevertheless, the content of the three subtests selected closely parallels the preskills listed by Calfee in his concept paper in reading and writing, and by Heller and Resnick in their mathematics paper.

The BSSI was developed in such a way so as to be highly appropriate for an evaluation of preacademic skills:

> During the spring of 1971, 50 first grade and kindergarten teachers in the greater Philadelphia area were asked to describe in detail the educational and behavioral characteristics of four or five children in their classes. They were instructed to choose at least two children who were 'ready' for school and two who were 'not ready'. They were then to specify the particular skills [tasks]that the pupils could or could not do that have prompted their selection. A wide variety of children were described. The sample included children from urban, suburban, and rural

residences and from both economically rich and depressed neighborhoods. Non-white pupils were included in approximately the same percentage as that in the community. An equal number of boys and girls were described. The most frequently mentioned skills designated by the teacher as being associated with ready and non-ready children were converted into 67 specific items. On the basis of face validity, each item was assigned to a different category of readiness: BASIC INFORMATION, SELF HELP, ORAL COMMUNICATION, HAND WRITING, READING READINESS, NUMBER READINESS, AND CLASSROOM BEHAVIOR. This collection of items comprised the first experimental form of BSSI. Only items that related directly to school performance and that were presumed to be teachable were included. For example, items related to the health of the child or to his home environment were excluded; important as these factors may be, there is little the teacher can do about them. (Goodman and Hammill, 1975)

This, and the second experimental form, were then tested and revised, and the final 84 item version of the BSSI completed.

The standardization sample of the BSSI consisted of 459 children, aged 4 to 6 years, half boys and half girls. Seventy percent of the sample were white; 22 percent were black; and 8 percent were Mexican-American. Half lived in urban areas, one-fourth in suburban and one-fourth in rural communities.

As stated earlier, the BSSI is an informal assessment device, intended to be used by classroom teachers. For use in the present evaluation, it was necessary to write complete item instructions and design materials so that administrative procedures would be consistent across testers. Certain items also had to be deleted or rewritten as they did not meet our testing demands. A description of each subtest follows.

Basic Information. The original subtest contains 11 items. Three items were deleted. "Can the child name any three animals, excluding dog and cat," was deleted as a child's knowledge cannot be assessed in a one or two session test. An item asking the child to differentiate betwen winter, summer and night time scenes was deleted as possibly cuturally biased. Finally, the item asking the child to tell the tester where he lived was eliminated as ambiguous to score. The remaining items all concern "basic information" or world knowledge, such as days of the week, coin, color, and shape names, etc.

Reading Readiness. We retained four of the original twelve items. All items concerning auditory discrimination (five items) were eliminated due to the difficulties in assessing correct responses among children who speak different languages or dialects. Three other items were deleted as they were difficult to assess in one session (i.e., determine the child's sight vocabulary) or difficult to score. The remaining items all concern skills known to relate to reading achievement, including knowledge of the alphabet, word matching and an awareness of proper left to right sequence.

Number Readiness. Of the original 11 items, 6 were retained. Four concern counting and/or numeration; one concerns first, second, third sequencing and one concerns verbal comparativeness (more/less, etc.). The eliminated items were deemed redundant with other items in the battery.

Preschool Interpersonal Problem Solving Test

The PIPS was the only measure of social skills we reviewed in depth. Examinations of other tests quickly showed that the PIPS is clearly the most well-designed and appropriate measure of social skills available. The test is appropriate in that social problem-solving ability which the PIPS measures

> "has perhaps the clearest relation to what is commonly thought of as social competence, and is integral to Bronson's definition of social executive skills, which she saw as 'the ability to control and direct oneself adequately and constructively in social situations and the ability to influence others effectively in socially approved ways.' (1975). It does not imply mere compliance, but rather the effective use of coping strategies. A member of the High/Scope PDC Advisory Panel, Charles Billings, stressed the importance of this category, emphasiing that the child's learning to control his or her environment and to influence others was an essential component of social competence and a necessary precursor to the development of social and political power, the knowledgeable use of which are vital for effective functioning in a democracy." (Love, Wacker, & Meece, 1975, p.23)

The PIPS was designed to test a child's ability to solve real life problems. The instrument is included here as it fulfills the need for a measure of problem solving and also provides us with some information on the child's social skills. Although the original test has two parts, one involving ways for a

child to obtain a toy from another child and the other concerning ways to avoid mother's anger when the child has broken something, only the child's toy situation was used here. The original test is far too long for our needs and for the attention span of the children in this investigation. Also, according to the Rand report, the first half of the test "is eminantly suited for assessing the exent to which Head Start increases prosocial behavior alternatives." (Raisen, et.al., 1974). Responses are scored according both to the number of solutions the child offers to a problem as well as the total number of different solution categories.

The difficulty with many measures of social skills (other than observational measures) is that they may actually assess the child's test taking skills or level of intellectual functioning, rather than social skills. An earlier study by Shure, Spivak and Jaeger (1971) found evidence that this is not true of the PIPS. They collected data on both the classroom behavior and PIPS performance of 62 four-year old Head Start children. They found that, regardless of sex and intellectual functioning, "those who are better adjusted in social behavior as measured are those who can conceptualize the greater variety of ways to handle interpersonal problems." (p.1799) Also, although the two behavioral groups (labeled "aberrant" and "nonabberant") differed significantly in the total number of solutions given on the PIPS, there was no difference between the two groups in mean IQ, as measured by the Peabody Picture Vocabulary Test.

References

Backe, W.L., Brush, L., Larson, J., & Mathis, A.L. An Approach Paper: Defining the Program Elements of the Head Start Demonstration on Basic Educational Skills. Cambridge, MA: Abt Associates, Inc., January, 1979.

Goodman, L., & Hammill, D.D. Basic School Skills Inventory. Chicago: Follett Publishing Company, 1975.

Kaufman, A.S., & Kaufman, N.L. Black-white differences at age 2 1/2-8 1/2 on the McCarthy Scales of Children's Abilities. Journal of School of Psychology, 1973, 11(3), 196-206.

Klein, J.W. Head Start Basic Educational Skills Project: A research and demonstration initiative. Phase I guidelines. Washington, D.C.; DHEW, 1978.

Love, J.M., Wacker, S., & Meece, J. A Process Evaluation of Project Developmental Continuity. Interim Report II, Part B. Ypsilanti, MI: High/Scope Foundation, June, 1975.

McCarthy, D. McCarthy Scales of Children's Abilities; New York: The Psychological Corporation, 1972.

Raizen, S. & Bobrow, S. Design for a National Evaluation of Social Competence in Head Start Children. Santa Monica: Rand Corporation, November, 1974.

Shure, M.B., & Spivak, G., The PIPS Test Manual. Philadelphia: Hahneman Medical College, 1974.

Shure, M.B., Spivak, G., & Jaeger, M. Problem-solving thinking and adjustment among disadvanaged preschool children. Child Development, 1971, 42, 1791-1803.

Wedell-Monnig, J., and McNeil, J. National Evaluation of Head Start Educational Services and Basic Educational Skills Initiative. Child Outcomes: Individual Testing Battery for the Fall 1979 Cohort. Durham, North Carolina: NTS Research Corporation, October, 1979.